Náhodná veličina

Michal Fusek

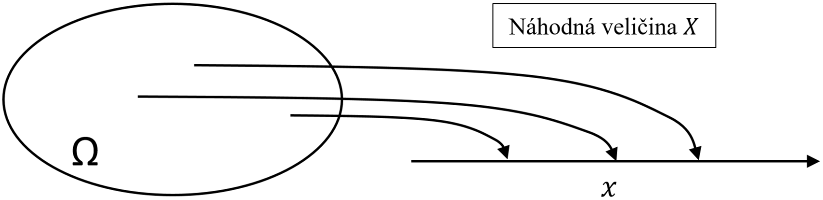
Ústav matematiky FEKT VUT, fusekmi@feec.vutbr.cz

Pravděpodobnost a statistika (IPT)

Obsah

Náhodná veličina

- Náhodný pokus, jehož výsledek jsme schopni číselně ohodnotit.
- Číselné ohodnocení výsledku náhodného pokusu nazveme náhodnou veličinou.
- Náhodná veličina se značí velkým písmenem, např. X, Y, Z.



Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Funkce $X\colon \Omega \to \mathbb{R}$ taková, že pro každé $x \in \mathbb{R}$ je $\{\omega \in \Omega\colon X(\omega) \le x\} \in \mathcal{A}$, se nazývá **náhodná veličina** na pravděpodobnostním prostoru $(\Omega, \mathcal{A}, P)$.

Příklad

Hodíme dvěma mincemi. Náhodná veličina X udává, kolikrát padl líc.

$$\Omega = \{RR, RL, LR, LL\}$$

Náhodná veličina X přiřazuje elementárním jevům číselné hodnoty:

$$X(RR) = 0, \quad X(RL) = 1, \quad X(LR) = 1, \quad X(LL) = 2.$$

Náhodný jev

- „líc padl právě jednou“ lze vyjádřit jako $[X = 1]$.
- „líc padl aspoň jednou“ lze vyjádřit jako $[X \geq 1]$.

Obecně zápisem $[X = x]$ rozumíme náhodný jev složený ze všech $\omega \in \Omega$, pro která je $X(\omega) = x$, tedy

$$[X = x] = \{\omega \in \Omega \colon X(\omega) = x\}.$$

Pravděpodobnost, že náhodná veličina X nabývá hodnoty x zapíšeme jako

$$P(X = x).$$

Podobně lze interpretovat $P(X < x)$, $P(X \geq x)$, atd.

Příklad

Hodíme dvěma mincemi. Náhodná veličina X udává, kolikrát padl líc.

$\Omega = \{RR, RL, LR, LL\}$

Obor hodnot náhodné veličiny X:

$X \in \{0, 1, 2\}$

Pravděpodobnost jednotlivých hodnot:

$P(X = 0) = 1/4, \qquad P(X = 1) = 2/4, \qquad P(X = 2) = 1/4$

Součet pravděpodobností je 1.

Distribuční funkce

- K popisu pravděpodobnostního chování náhodné veličiny používáme různé funkce a charakteristiky.

Distribuční funkce náhodné veličiny X je funkce $F\colon \mathbb{R} \to \langle 0, 1\rangle$, která je definována jako

$$F(x) = P(X \leq x).$$

Distribuční funkce $F(x)$ náhodné veličiny X přiřazuje každému reálnému číslu x pravděpodobnost, že náhodná veličina X nabude hodnoty menší nebo rovné číslu x.

Vlastnosti distribuční funkce

- $0 \leq F(x) \leq 1$ pro $\forall\, x \in \mathbb{R}$
- $F(x)$ je neklesající a zprava spojitá funkce.
- $\lim\limits_{x\to-\infty} F(x) = 0,\ \lim\limits_{x\to\infty} F(x) = 1$
- $P(a < X \leq b) = F(b) - F(a)$ pro každé $a, b \in \mathbb{R}$, $a < b$
- $P(X = x) = F(x) - \lim\limits_{t\to x^-} F(t)$
- F má nejvýše spočetně mnoho bodů nespojitosti

Rozlišujeme dva základní typy náhodných veličin:

- **diskrétní**
- **spojitá**

Diskrétní náhodná veličina

- Má **konečný** nebo **spočetný** obor hodnot M.
- Existuje **nezáporná** funkce $p(x)$, pro kterou platí

$$\sum_{x \in M} p(x) = 1.$$

- Distribuční funkci $F(x)$ lze vyjádřit ve tvaru

$$F(x) = P(X \leq x) = \sum_{t \in (-\infty, x\rangle \cap M} p(t).$$

- Distribuční funkce je schodovitá.

Pravděpodobnostní funkce

Pravděpodobnostní funkce diskrétní náhodné veličiny X je funkce $p\colon \mathbb{R} \to \langle 0, 1\rangle$, která je definována jako

$$p(x) = P(X = x)$$

pro všechna x z oboru hodnot náhodné veličiny X a $p(x) = 0$ jinak.

Nechť X je diskrétní náhodná veličina. Pak

- $p(x) \geq 0$ pro $\forall\, x \in \mathbb{R}$ a $\displaystyle\sum_{x \in M} p(x) = 1$
- $F(x) = P(X \leq x) = \displaystyle\sum_{t \leq x} p(t)$ pro $\forall\, x \in \mathbb{R}$
- $P(X \in A) = \displaystyle\sum_{x \in M \cap A} p(x)$
- $p(x) = F(x) - \displaystyle\lim_{t \to x^-} F(t)$ pro $\forall\, x \in \mathbb{R}$

Příklad

V televizní soutěži dostává soutěžící postupně otázky, u kterých jsou na výběr vždy tři možné odpovědi. Jestliže soutěžící odpoví správně, dostane další otázku. Jestliže odpoví špatně, soutěž končí. Nanejvýš však může dostat čtyři otázky, pak soutěž končí každopádně. Náhodná veličina X udává, na kolik otázek soutěžící správně odpoví, jestliže všechny odpovědi pouze náhodně tipuje.

a) Popište veličinu X pomocí pravděpodobnostní a distribuční funkce a nakreslete grafy obou funkcí.

b) Určete pravděpodobnost, že soutěžící správně odpoví alespoň na dvě otázky.

Řešení:

X...počet správných odpovědí

$X \in \{0, 1, 2, 3, 4\}$

Pravděpodobnost, že soutěžící nezodpoví žádnou otázku správně:

$$p(0) = P(X = 0) = \frac{2}{3}$$

Podobně:

$$p(1) = P(X = 1) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

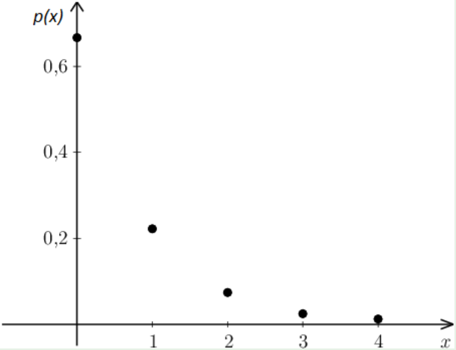
$$p(2) = P(X = 2) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{27}$$

$$p(3) = P(X = 3) = \left(\frac{1}{3}\right)^3 \cdot \frac{2}{3} = \frac{2}{81}$$

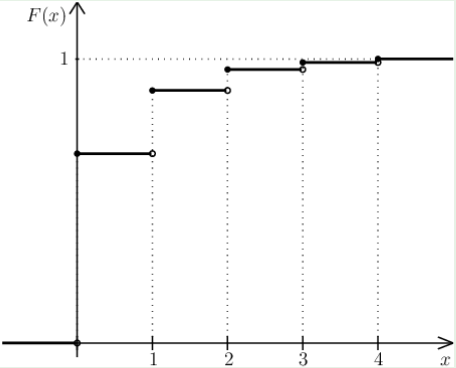
$$p(4) = P(X = 4) = \left(\frac{1}{3}\right)^4 = \frac{1}{81}$$

$$\sum_{i=0}^{4} p(i) = 1$$

x	0	1	2	3	4	jinak
$p(x)$	$\frac{2}{3}$	$\frac{2}{9}$	$\frac{2}{27}$	$\frac{2}{81}$	$\frac{1}{81}$	0



$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ \frac{2}{3} & \text{pro } x \in \langle 0, 1) , \\ \frac{8}{9} & \text{pro } x \in \langle 1, 2) , \\ \frac{26}{27} & \text{pro } x \in \langle 2, 3) , \\ \frac{80}{81} & \text{pro } x \in \langle 3, 4) , \\ 1 & \text{pro } x \in \langle 4, \infty) . \end{cases}$$



$$P(X \geq 2) = p(2) + p(3) + p(4) = \frac{1}{9}$$

Charakteristiky náhodné veličiny

Střední hodnotu diskrétní náhodné veličiny X s oborem hodnot M označíme $\mathrm{E}(X)$ a definujeme ji vztahem

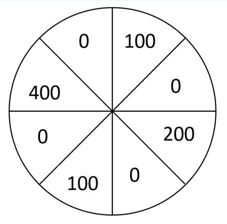
$$\mathrm{E}(X) = \sum_{x \in M} x \cdot p(x).$$

Nechť X, Y jsou náhodné veličiny a dále $a, b \in \mathbb{R}$. Pak platí:

- $\mathrm{E}(a) = a$
- $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$
- $\mathrm{E}(X \pm Y) = \mathrm{E}(X) \pm \mathrm{E}(Y)$

Příklad

Na pouti mají kolo štěstí, kde každé pole definuje výhru v Kč. Jedno zatočení stojí 100 Kč. Vyplatí se takovou hru hrát (jaký je očekávaný zisk)?



Řešení:

X...částka vytočená na kole v jedné hře

$X \in \{0, 100, 200, 400\}$

x	0	100	200	400
$p(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$

$$\mathrm{E}(X) = \sum_x x \cdot p(x) = 100$$

$Y \ldots$ zisk v jedné hře

$Y = X - 100$

$Y \in \{-100, 0, 100, 300\}$

$$\mathrm{E}(Y) = \mathrm{E}(X - 100) = \mathrm{E}(X) - 100 = 0$$

Rozptyl diskrétní náhodné veličiny X s oborem hodnot M označíme $\mathrm{D}(X)$ a definujeme jej vztahem

$$\mathrm{D}(X) = \mathrm{E}\left[X - \mathrm{E}(X)\right]^2 = \sum_{x \in M} \left[x - \mathrm{E}(X)\right]^2 \cdot p(x).$$

Platí

$$\mathrm{D}(X) = \mathrm{E}(X^2) - \left[\mathrm{E}(X)\right]^2,$$

kde

$$\mathrm{E}(X^2) = \sum_{x \in M} x^2 \cdot p(x).$$

Nechť X, Y jsou náhodné veličiny a dále $a, b \in \mathbb{R}$. Pak platí:

- $\mathrm{D}(X) \geq 0$
- $\mathrm{D}(aX + b) = a^2\,\mathrm{D}(X)$
- $\mathrm{D}(X \pm Y) = \mathrm{D}(X) + \mathrm{D}(Y)$ pro nezávislé náhodné veličiny X, Y

Číslo

$$\sigma(X) = \sqrt{\mathrm{D}(X)}$$

nazveme **směrodatnou odchylkou** náhodné veličiny X.

Obecné a centrální momenty

Nechť X je náhodná veličina definovaná na $(\Omega, \mathcal{A}, P)$. Pak číslo

- $\mu'_k = \mathrm{E}\left(X^k\right)$ nazveme k-tým **obecným momentem**,
- $\mu_k = \mathrm{E}\left[X - \mathrm{E}(X)\right]^k$ nazveme k-tým **centrálním momentem**

náhodné veličiny X, pokud uvedené střední hodnoty pro $k = 1, 2, \ldots$ existují.

Nechť pro $n \in \mathbb{N}$ existuje $\mu'_n = \mathrm{E}(X^n) < \infty$. Potom existuje i $\mu'_k = \mathrm{E}\left(X^k\right) < \infty$ pro libovolné $k \leq n$.

Alternativní rozdělení

- Některé modely rozdělení pravděpodobnosti mají vlastní názvy.
- Známe jejich pravděpodobnostní/distribuční funkci, střední hodnotu, rozptyl.

Provádíme pokus, u kterého úspěch nastane s pravděpodobností π, $\pi \in (0,1)$, a neúspěch s pravděpodobností $1-\pi$. Náhodná veličina X, která udává, zda úspěch nastal ($X=1$), nebo nenastal ($X=0$), má **alternativní rozdělení** pravděpodobnosti, což zapíšeme jako

$$X \sim A(\pi).$$

Dále

$$p(x) = \begin{cases} \pi^x (1-\pi)^{1-x} & \text{pro } x = 0, 1, \\ 0 & \text{jinak.} \end{cases}$$

$\mathrm{E}(X) = \pi$, $\mathrm{D}(X) = \pi(1-\pi)$.

Příklad

Střelec vystřelí do terče, pravděpodobnost zásahu je 0,8. Náhodná veličina X udává, zda se trefil, nebo netrefil. Najděte pravděpodobnostní a distribuční funkci, střední hodnotu a rozptyl náhodné veličiny X.

Řešení:

X...počet zásahů

$\pi = 0{,}8$

x	0	1	jinak
$p(x)$	0,2	0,8	0

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ 0{,}2 & \text{pro } x \in \langle 0, 1), \\ 1 & \text{pro } x \in \langle 1, \infty). \end{cases}$$

$$\mathrm{E}(X) = 0{,}8$$

$$\mathrm{D}(X) = 0{,}16$$

Binomické rozdělení

Postupně n-krát nezávisle opakujeme pokus, u kterého úspěch nastává s pravděpodobností π. Náhodná veličina X udávající kolikrát v těchto n pokusech nastal úspěch má **binomické rozdělení** pravděpodobnosti s parametry n a π, píšeme

$$X \sim Bi(n, \pi).$$

Její pravděpodobnostní funkce je

$$p(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, \quad x = 0, 1, \dots, n.$$

Střední hodnota a rozptyl jsou

$$\begin{aligned} \mathrm{E}(X) &= n\pi, \\ \mathrm{D}(X) &= n\pi(1-\pi). \end{aligned}$$

Mějme posloupnost nezávislých náhodných veličin $X_1, \ldots, X_n$, kde

$$X_i \sim A(\pi), \quad i = 1, \ldots, n.$$

Jestliže

$$Y = \sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n,$$

pak

$$Y \sim Bi(n, \pi).$$

Alternativní rozdělení je tedy speciálním případem binomického rozdělení.

Příklad

Hodíme čtyřikrát kostkou. Náhodná veličina X udává, kolikrát padla šestka. Najděte pravděpodobnostní a distribuční funkci, střední hodnotu a rozptyl náhodné veličiny X.

Řešení:

X...počet šestek

$X \in \{0, 1, 2, 3, 4\}$

$\pi = 1/6$

$$p(0) = \left(\frac{5}{6}\right)^4 = \frac{625}{1296}$$

$$p(1) = 4 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 = \frac{500}{1296}$$

$$p(2) = 6 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2 = \frac{150}{1296}$$

$$p(3) = 4 \cdot \left(\frac{1}{6}\right)^3 \cdot \frac{5}{6} = \frac{20}{1296}$$

$$p(4) = \left(\frac{1}{6}\right)^4 = \frac{1}{1296}$$

x	0	1	2	3	4
$p(x)$	$\frac{625}{1296}$	$\frac{500}{1296}$	$\frac{150}{1296}$	$\frac{20}{1296}$	$\frac{1}{1296}$

$\mathrm{E}(X) = \frac{2}{3}$

$\mathrm{D}(X) = \frac{5}{9}$

$$p(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}, \quad x = 0, 1, \ldots, n$$

$$p(x) = \binom{4}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{4-x}, \quad x = 0, 1, 2, 3, 4$$

x	0	1	2	3	4
$p(x)$	$\frac{625}{1296}$	$\frac{500}{1296}$	$\frac{150}{1296}$	$\frac{20}{1296}$	$\frac{1}{1296}$

$$F(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0), \\ \frac{625}{1296} & \text{pro } x \in \langle 0, 1), \\ \frac{1125}{1296} & \text{pro } x \in \langle 1, 2), \\ \frac{1275}{1296} & \text{pro } x \in \langle 2, 3), \\ \frac{1295}{1296} & \text{pro } x \in \langle 3, 4), \\ 1 & \text{pro } x \in \langle 4, \infty). \end{cases}$$

Hypergeometrické rozdělení

Máme množinu o N prvcích, z nichž M má sledovanou vlastnost. Náhodně vybereme (bez vracení) n prvků. Náhodná veličina X udávající, kolik z vybraných n prvků má sledovanou vlastnost, má **hypergeometrické rozdělení** pravděpodobnosti s parametry N, M a n, píšeme

$$X \sim Hg(N, M, n).$$

Její pravděpodobnostní funkce je

$$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = \max\{0, n-(N-M)\}, \dots, \min\{n, M\}.$$

Střední hodnota a rozptyl jsou

$$\mathrm{E}(X) = n\frac{M}{N}, \qquad \mathrm{D}(X) = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}.$$

Příklad

Máme 12 výrobků, z nichž 4 jsou vadné. Náhodně vybereme 3 výrobky. Náhodná veličina X udává, kolik výrobků z vybrané trojice je vadných. Najděte pravděpodobnostní a distribuční funkci, střední hodnotu a rozptyl náhodné veličiny X.

Řešení:

X...počet vadných ze 3 vybraných

$X \in \{0, 1, 2, 3\}$

$$p(0) = \frac{\binom{4}{0}\binom{8}{3}}{\binom{12}{3}} = \frac{14}{55}$$

$$p(1) = \frac{\binom{4}{1}\binom{8}{2}}{\binom{12}{3}} = \frac{28}{55}$$

$$p(2) = \frac{\binom{4}{2}\binom{8}{1}}{\binom{12}{3}} = \frac{12}{55}$$

$$p(3) = \frac{\binom{4}{3}\binom{8}{0}}{\binom{12}{3}} = \frac{1}{55}$$

x	0	1	2	3
$p(x)$	$\frac{14}{55}$	$\frac{28}{55}$	$\frac{12}{55}$	$\frac{1}{55}$

$$\mathrm{E}(X) = 1$$

$$\mathrm{D}(X) = \tfrac{6}{11}$$

$$p(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, \quad x = \max\{0, n-(N-M)\}, \dots, \min\{n, M\}$$

$$p(x) = \frac{\binom{4}{x}\binom{8}{3-x}}{\binom{12}{3}}, \quad x = 0, 1, 2, 3$$

x	0	1	2	3
$p(x)$	$\frac{14}{55}$	$\frac{28}{55}$	$\frac{12}{55}$	$\frac{1}{55}$

$$F(x)=\begin{cases} 0 & \text{pro } x\in(-\infty,0), \\ \frac{14}{55} & \text{pro } x\in\langle 0,1)\,, \\ \frac{42}{55} & \text{pro } x\in\langle 1,2)\,, \\ \frac{54}{55} & \text{pro } x\in\langle 2,3)\,, \\ 1 & \text{pro } x\in\langle 3,\infty)\,. \end{cases}$$

Poissonovo rozdělení

Budeme zkoumat výskyt událostí:

- příchod zákazníka do fronty,
- příjezd automobilu na parkoviště,
- dopravní nehoda v určitém úseku dálnice.

Podmínky:

- V jednom okamžiku může nastat nanejvýš jedna událost (tedy nemohou nastat dvě zcela současně).
- Události přicházejí nezávisle na sobě (počty vzniklých událostí v disjunktních časových intervalech jsou nezávislé).
- Pravděpodobnost, že událost nastane v intervalu $(t, t+h)$, závisí na h (délce intervalu), ale nikoli na t (umístění intervalu na časové ose).

Poissonovo rozdělení

Náhodná veličina X, která udává počet událostí za jednotku času, když víme, že průměrně nastává λ událostí za jednotku času, má **Poissonovo rozdělení** pravděpodobnosti s parametrem λ, píšeme

$$X \sim Po(\lambda).$$

Její pravděpodobnostní funkce je

$$p(x) = \frac{\lambda^x}{x!}\,\mathrm{e}^{-\lambda}, \quad x = 0, 1, 2, \ldots$$

Střední hodnota a rozptyl jsou

$$\mathrm{E}(X) = \lambda,$$
$$\mathrm{D}(X) = \lambda.$$

Příklad

Zdravotnický úřad shromažďuje údaje o nově narozených dětech. Průměrně se každé dvě hodiny narodí další dítě. Určete pravděpodobnost, že se

a) v daném dnu nenarodí žádné dítě.

b) za 3 hodiny narodí aspoň 4 děti.

Řešení:

a) X...počet narozených dětí za den

$X \in \{0, 1, 2, \dots\}$

λ...průměrný počet dětí narozených za den $\Rightarrow \quad \lambda = 12$

$$p(0) = \frac{12^0}{0!}\, \mathrm{e}^{-12} \doteq 0{,}00000614$$

b) X...počet narozených dětí za 3 hodiny

$X \in \{0, 1, 2, \dots\}$

λ...průměrný počet dětí narozených za 3 hodiny $\quad \Rightarrow \quad \lambda = 1{,}5$

$$P(X \geq 4) = 1 - P(X < 4) \doteq 0{,}0656$$

Geometrické rozdělení

Postupně nezávisle opakujeme pokus, u kterého úspěch nastává s pravděpodobností π. Náhodná veličina X udávající počet úspěchů před prvním neúspěchem má **geometrické rozdělení** pravděpodobnosti s parametrem π, píšeme

$$X \sim Ge(\pi).$$

Její pravděpodobnostní funkce je

$$p(x) = \pi^x(1-\pi), \quad x = 0, 1, \dots$$

Střední hodnota a rozptyl jsou

$$\mathrm{E}(X) = \frac{\pi}{1-\pi},$$
$$\mathrm{D}(X) = \frac{\pi}{(1-\pi)^2}.$$

- Často je výhodnější pracovat raději s veličinou $Y = X + 1$, která udává, v kolikátém pokusu nastane neúspěch.
- Pravděpodobnostní funkce je pak tvaru:

$$p(y) = \pi^{y-1}(1 - \pi), \quad y = 1, 2, \ldots$$

Příklad

Studenti hrají před fakultou Frisbee. Pepovi to moc nejde a s pravděpodobností $1/5$ trefí náhodného kolemjdoucího. Nechť X je náhodná veličina označující počet hodů, než Pepa trefí kolemjdoucího. Jaká je pravděpodobnost, že

a) omylem trefí kolemjdoucího až v 5. hodu?
b) netrefí žádného kolemjdoucího v prvních 10 hodech?
c) nebude trvat více než 7 hodů, než někoho trefí?

Řešení:

X...počet hodů, než někoho trefí

$X \in \{0, 1, 2, \dots\}$

π...pravděpodobnost, že netrefí

$$\pi = \frac{4}{5}$$

$Y = X + 1$...v kolikátém hodu někoho trefí

$Y \in \{1, 2, 3, \dots\}$

a) omylem trefí kolemjdoucího až v 5. hodu

$$P(Y = 5) = p(5) = \left(\frac{4}{5}\right)^4 \cdot \frac{1}{5} \doteq 0{,}082$$

b) netrefí žádného kolemjdoucího v prvních 10 hodech

$$P(Y > 10) = \left(\frac{4}{5}\right)^{10} \doteq 0{,}107$$

c) nebude trvat více než 7 hodů, než někoho trefí

$$P(Y \leq 7) = 1 - P(Y > 7) = 1 - \left(\frac{4}{5}\right)^7 \doteq 0{,}790$$

Spojitá náhodná veličina

- Má **nespočetný** obor hodnot.
- Existuje **nezáporná** funkce $f(x)$, pro kterou $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$.
- Distribuční funkci $F(x)$ lze vyjádřit ve tvaru

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

- Distribuční funkce je spojitá.

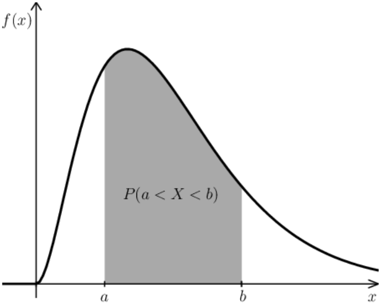
Hustota pravděpodobnosti spojité náhodné veličiny X je nezáporná funkce $f(x)$ taková, že

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t, \quad x \in \mathbb{R}.$$

Platí:

- $f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} = F'(x)$ pro všechna x, kde derivace existuje.
- $$\begin{aligned} P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b) \\ = F(b) - F(a) \\ = \int_a^b f(x)\,\mathrm{d}x \end{aligned}$$

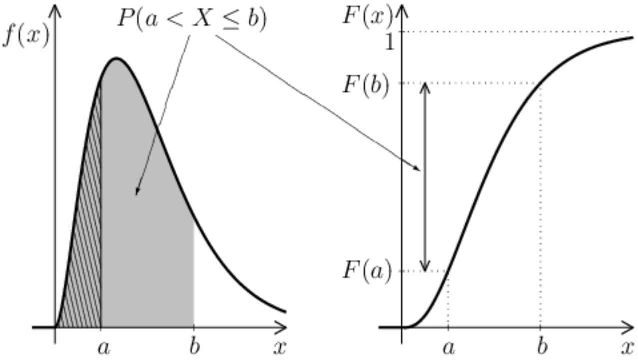
$\Rightarrow \; P(X = x) = 0$



Vztah mezi hustotou a distribuční funkcí

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$$

$$P(a < X \leq b) = F(b) - F(a) = \int_{a}^{b} f(x)\,\mathrm{d}x$$



Střední hodnota a rozptyl

Střední hodnotu spojité náhodné veličiny X označíme $\mathrm{E}(X)$ a definujeme ji vztahem

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x.$$

Rozptyl spojité náhodné veličiny X označíme $\mathrm{D}(X)$ a definujeme jej vztahem

$$\mathrm{D}(X) = \mathrm{E}\,[X - \mathrm{E}(X)]^2 = \int_{-\infty}^{\infty} [x - \mathrm{E}(X)]^2 \cdot f(x)\,\mathrm{d}x.$$

Platí

$$\mathrm{D}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2\,,$$

kde

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,\mathrm{d}x.$$

Kvantil

Jestliže X je náhodná veličina, jejíž distribuční funkce F je prostá, a $\alpha \in (0,1)$, pak α-**kvantilem** náhodné veličiny X nazveme to číslo $x_\alpha \in \mathbb{R}$, pro které platí

$$P(X \leq x_\alpha) = F(x_\alpha) = \alpha \quad \text{neboli} \quad x_\alpha = F^{-1}(\alpha).$$

Když distribuční funkce není prostá, definujeme α-kvantil jako to číslo $x_\alpha \in \mathbb{R}$, pro které platí

$$P(X \leq x_\alpha) \geq \alpha \quad \text{a současně} \quad P(X \geq x_\alpha) \geq 1 - \alpha.$$

Speciálně 0,5-kvantil se nazývá **medián** náhodné veličiny X.

Příklad

Náhodná veličina X má rozdělení dáno funkcí

$$f(x) = \begin{cases} c - 2x & 0 < x < 1, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete c tak, aby funkce f byla hustotou a nakreslete graf.
b) Určete distribuční funkci a nakreslete graf.
c) Určete $P(X \geq 0{,}5)$, $P(0 < X \leq 0{,}75)$, $P(X = 0{,}25)$.
c) Určete střední hodnotu a rozptyl veličiny X.

Řešení:

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

$$\int_0^1 c - 2x\,\mathrm{d}x = c - 1 \qquad \Rightarrow \qquad c = 2$$

$$f(x) = \begin{cases} 2 - 2x & 0 < x < 1, \\ 0 & \text{jinak.} \end{cases}$$

$$x \leq 0: \quad F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} 0\,\mathrm{d}t = 0$$

$$0 < x < 1: \quad F(x) = \int_{-\infty}^{0} 0\,\mathrm{d}t + \int_{0}^{x} 2 - 2t\,\mathrm{d}t = 2x - x^2$$

$$x \geq 1: \quad F(x) = \int_{-\infty}^{0} 0\,\mathrm{d}t + \int_{0}^{1} 2 - 2t\,\mathrm{d}t + \int_{1}^{x} 0\,\mathrm{d}t = 1$$

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ 2x - x^2 & \text{pro } 0 < x < 1, \\ 1 & \text{pro } x \geq 1. \end{cases}$$

$$P(X \geq 0{,}5) = 1 - P(X < 0{,}5) = 1 - F(0{,}5) = 0{,}25$$

$$P(0 < X \leq 0{,}75) = F(0{,}75) - F(0) = 0{,}9375$$

$$P(X = 0{,}25) = 0$$

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x = \int_0^1 x(2-2x)\,\mathrm{d}x = \frac{1}{3}$$

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,\mathrm{d}x = \int_0^1 x^2(2-2x)\,\mathrm{d}x = \frac{1}{6}$$

$$\mathrm{D}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{1}{18}$$

Příklad

Náhodná veličina X má rozdělení dáno funkcí

$$f(x) = \begin{cases} \frac{2}{5} & 0 < x \leq 1, \\ cx & 1 < x < 2, \\ 0 & \text{jinak.} \end{cases}$$

a) Určete c tak, aby funkce f byla hustotou a nakreslete graf.

b) Určete distribuční funkci a nakreslete graf.

c) Určete $P(X \geq 1{,}5)$.

c) Určete střední hodnotu a rozptyl veličiny X.

Řešení:

$$\int_0^1 \frac{2}{5}\,\mathrm{d}x + \int_1^2 cx\,\mathrm{d}x = \frac{2}{5} + \frac{3}{2}c \qquad \Rightarrow \qquad c = \frac{2}{5}$$

$$f(x) = \begin{cases} \frac{2}{5} & 0 < x \leq 1, \\ \frac{2}{5}x & 1 < x < 2, \\ 0 & \text{jinak.} \end{cases}$$

$$x \leq 0: \quad F(x) = 0$$

$$0 < x \leq 1: \quad F(x) = \int_0^x \frac{2}{5}\,\mathrm{d}t = \frac{2}{5}x$$

$$1 < x < 2: \quad F(x) = \int_0^1 \frac{2}{5}\,\mathrm{d}t + \int_1^x \frac{2}{5}t\,\mathrm{d}t = \frac{x^2}{5} + \frac{1}{5}$$

$$x \geq 2: \quad F(x) = \int_0^1 \frac{2}{5}\,\mathrm{d}t + \int_1^2 \frac{2}{5}t\,\mathrm{d}t = 1$$

$$F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \frac{2}{5}x & 0 < x \leq 1, \\ \frac{x^2}{5} + \frac{1}{5} & 1 < x < 2, \\ 1 & \text{pro } x \geq 2. \end{cases}$$

$$P(X \geq 1{,}5) = 1 - P(X < 1{,}5) = 1 - F(1{,}5) = \frac{7}{20}$$

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x = \int_0^1 \frac{2}{5}x\,\mathrm{d}x + \int_1^2 \frac{2}{5}x^2\,\mathrm{d}x = \frac{17}{15}$$

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,\mathrm{d}x = \int_0^1 \frac{2}{5}x^2\,\mathrm{d}x + \int_1^2 \frac{2}{5}x^3\,\mathrm{d}x = \frac{49}{30}$$

$$\mathrm{D}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{157}{450}$$

Příklad

Životnost určitého typu součástek (měřená ve stovkách hodin) je náhodná veličina X s hustotou

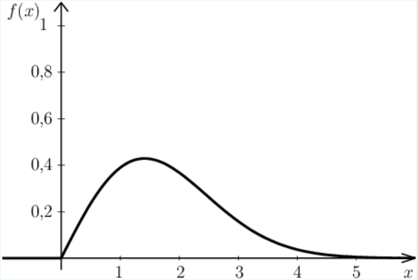
$$f(x) = \begin{cases} \frac{x}{2}\,\mathrm{e}^{-\frac{x^2}{4}} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0. \end{cases}$$

a) Určete distribuční funkci veličiny X.
b) Určete pravděpodobnost, že součástka vydrží 100 až 300 hodin.
c) Určete pravděpodobnost, že součástka vydrží nanejvýš 200 hodin.
d) Pod jakou hranicí bude životnost s pravděpodobností 0,75?

Řešení:

X...životnost součástky

$X \in (0, \infty)$

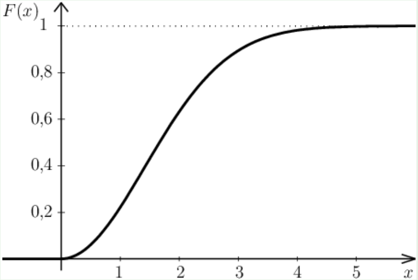


$$f(x) = \begin{cases} \frac{x}{2}\,\mathrm{e}^{-\frac{x^2}{4}} & \text{pro } x > 0 \\ 0 & \text{pro } x \le 0 \end{cases}$$

$$x \le 0: \quad F(x) = 0$$

$$x > 0: \quad F(x) = \int_0^x \frac{t}{2}\,\mathrm{e}^{-\frac{t^2}{4}}\,\mathrm{d}t = \left[-\,\mathrm{e}^{-\frac{t^2}{4}}\right]_0^x = 1 - \mathrm{e}^{-\frac{x^2}{4}}$$

$$F(x) = \begin{cases} 1 - e^{-\frac{x^2}{4}} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0. \end{cases}$$



$$P(1 \leq X \leq 3) = F(3) - F(1) \doteq 0,6734$$

$$P(X \leq 2) = F(2) \doteq 0,6321$$

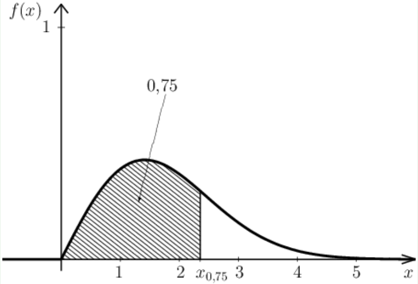
Pod jakou hranicí bude životnost s pravděpodobností 0,75?

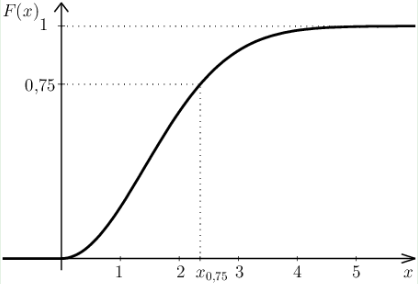
$$P(X \leq x_{0,75}) = 0{,}75$$

$$F(x_{0,75}) = 0{,}75$$

$$1 - \mathrm{e}^{-\frac{x_{0,75}^2}{4}} = 0{,}75$$

$$x_{0,75} \doteq 2{,}35 \quad \text{(235 hodin)}$$





Rovnoměrné rozdělení

Náhodná veličina X s **rovnoměrným rozdělením** na intervalu $\langle a, b\rangle$,

$$X \sim Ro(a, b),$$

má hustotu f a distribuční funkci F

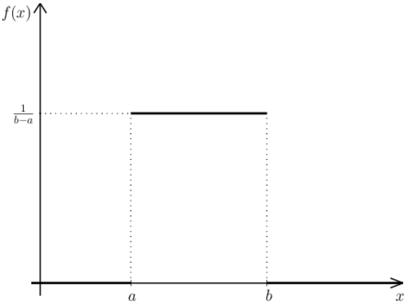
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in \langle a, b\rangle\,, \\ 0 & \text{jinak,} \end{cases} \qquad F(x) = \begin{cases} 0 & \text{pro } x < a, \\ \frac{x-a}{b-a} & \text{pro } x \in \langle a, b\rangle\,, \\ 1 & \text{pro } x > b. \end{cases}$$

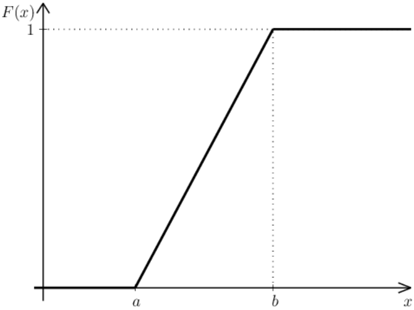
Střední hodnota a rozptyl jsou

$$\mathrm{E}(X) = \frac{a+b}{2}, \qquad \mathrm{D}(X) = \frac{(b-a)^2}{12}.$$

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in \langle a, b \rangle \\ 0 & \text{jinak} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{pro } x < a \\ \frac{x-a}{b-a} & \text{pro } x \in \langle a, b \rangle \\ 1 & \text{pro } x > b \end{cases}$$





Příklad

Pepa jezdí do práce šalinou, která jezdí v šestiminutových intervalech. Na zastávku přijde naprosto náhodně a čeká, až pojede šalina. Náhodná veličina X udává dobu čekání.

a) Popište veličinu X pomocí hustoty a distribuční funkce.

b) Vypočtěte pravděpodobnost, že bude čekat déle než 4 minuty.

c) Vypočtěte střední hodnotu a rozptyl náhodné veličiny X.

Řešení:

X...doba čekání

$$f(x) = \begin{cases} \frac{1}{6} & \text{pro } x \in (0,6) \\ 0 & \text{jinak} \end{cases} \qquad F(x) = \begin{cases} 0 & \text{pro } x \leq 0, \\ \frac{x}{6} & \text{pro } 0 < x < 6, \\ 1 & \text{pro } x \geq 6. \end{cases}$$

$$P(X > 4) = 1 - F(4) = \frac{1}{3}$$

$$= \int_4^\infty f(x)\mathrm{d}x = \int_4^6 \frac{1}{6}\mathrm{d}x = \frac{1}{3}$$

$$X \sim Ro(a,b) \;\Rightarrow\; X \sim Ro(0,6)$$

$$\mathrm{E}(X) = \frac{a+b}{2} = 3$$

$$\mathrm{D}(X) = \frac{(b-a)^2}{12} = 3$$

Exponenciální rozdělení

Nechť platí stejné předpoklady jako u Poissonova rozdělení. Náhodná veličina X, která udává dobu mezi dvěma výskyty určité události, když víme, že průměrně nastává λ událostí za jednotku času, má **exponenciální rozdělení** pravděpodobnosti s parametrem λ, píšeme

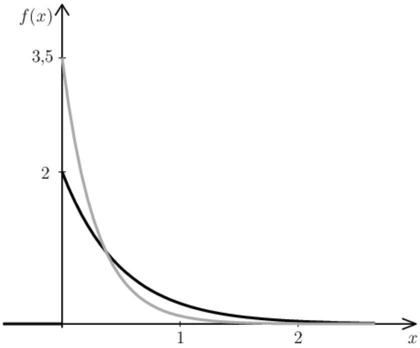
$$X \sim Exp(\lambda).$$

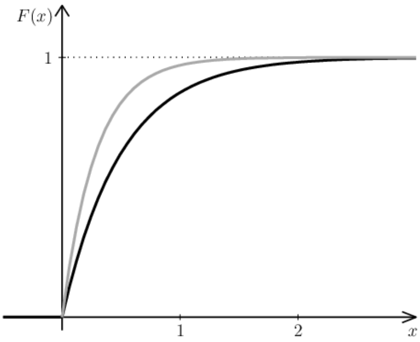
Hustota f a distribuční funkce F jsou

$$f(x) = \begin{cases} \lambda\,\mathrm{e}^{-\lambda x} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0, \end{cases} \qquad F(x) = \begin{cases} 1 - \mathrm{e}^{-\lambda x} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0. \end{cases}$$

Střední hodnota a rozptyl jsou

$$\mathrm{E}(X) = \frac{1}{\lambda}, \qquad \mathrm{D}(X) = \frac{1}{\lambda^2}.$$





$\lambda = 2$ (černě), $\lambda = 3{,}5$ (šedě)

Příklad

Přístroj má poruchu v průměru jednou za 2000 hodin. Veličina X představující dobu čekání na poruchu má exponenciální rozdělení.

a) Jaká je pravděpodobnost, že přístroj bude mít poruchu dříve než za 1000 hodin?
b) Určete dobu T_1 tak, aby pravděpodobnost, že přístroj bude pracovat delší dobu než T_1, byla 0,9.
c) Do jaké doby T_2 se přístroj pokazí s pravděpodobností 0,9?
d) Přístroj už pracuje bez poruchy 1000 hodin. Jaká je pravděpodobnost, že vydrží pracovat ještě alespoň 2000 hodin?

Řešení:

X...doba čekání na poruchu

$X \in (0, \infty)$

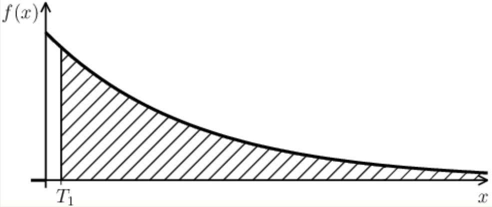
$$\left.\begin{array}{l} \mathrm{E}(X) = 2000 \\ \mathrm{E}(X) = \frac{1}{\lambda} \end{array}\right\} \Rightarrow \lambda = \frac{1}{2000}$$

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{2000}} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0. \end{cases}$$

a) Přístroj bude mít poruchu dříve než za 1000 hodin:

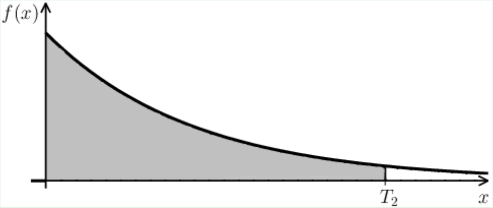
$$P(X < 1000) = F(1000) = 1 - e^{-\frac{1}{2}} \doteq 0{,}3934$$

b) Přístroj bude pracovat déle než T_1 s pravděpodobností 0,9:



$$\begin{aligned} P(X > T_1) &= 0{,}9 \\ 1 - P(X \leq T_1) &= 0{,}9 \\ 1 - F(T_1) &= 0{,}9 \\ e^{-\frac{T_1}{2000}} &= 0{,}9 \\ T_1 &\doteq 210{,}72 \end{aligned}$$

c) Do jaké doby T_2 se přístroj pokazí s pravděpodobností 0,9:



$$P(X < T_2) = 0{,}9$$

$$F(T_2) = 0{,}9$$

$$1 - \mathrm{e}^{-\frac{T_2}{2000}} = 0{,}9$$

$$T_2 \doteq 4605{,}17$$

d) Pracuje už 1000 hodin a vydrží ještě aspoň 2000 hodin:

$$P(X > 3000 | X > 1000) = \frac{P(X > 3000, X > 1000)}{P(X > 1000)} = \frac{P(X > 3000)}{P(X > 1000)}$$

$$= \frac{1 - F(3000)}{1 - F(1000)} = \mathrm{e}^{-1} \doteq 0{,}368$$

$$P(X > 2000) = 1 - F(2000) = \mathrm{e}^{-1}$$

Veličina s exponenciálním rozdělením nemá paměť - nezáleží na tom, jak dlouho už přístroj pracoval:

$$P(X > a + t | X > a) = P(X > t)$$

Modelování životnosti součástek, které nepodléhají opotřebení (např. elektronické a magnetické spínače, indukční vazební prvky).
Jinak spíš Weibullovo rozdělení.

Normální rozdělení

- Jedno z nejdůležitějších rozdělení.
- Použitelné tam, kde je kolísání náhodné veličiny způsobeno součtem velkého počtu nepatrných a vzájemně nezávislých vlivů (náhodné chyby při měřeních).
- Lze jím aproximovat (CLV) řadu dalších rozdělení.

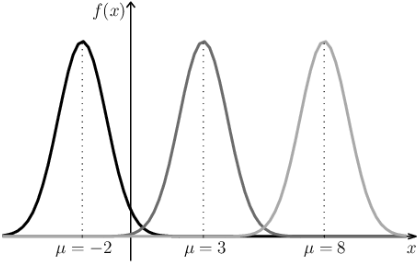
Náhodná veličina X s **normálním rozdělením** se střední hodnotou μ a rozptylem σ^2 má hustotu a distribuční funkci

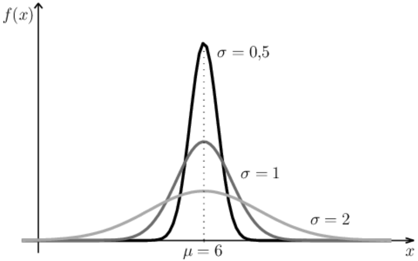
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t, \quad x \in \mathbb{R}.$$

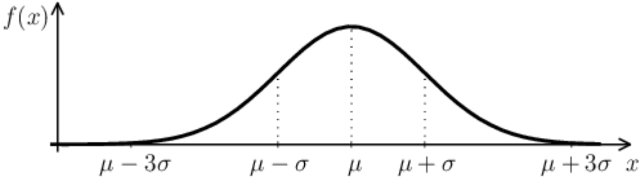
Zapisujeme

$$X \sim N(\mu, \sigma^2).$$





Mezi $\mu \pm 3\sigma$ leží 99,7 % hodnot:



Standardizované normální rozdělení

Normální rozdělení s parametry $\mu = 0$, $\sigma^2 = 1$ se nazývá **standardizované** (normované) normální rozdělení. Náhodnou veličinu s tímto rozdělením označíme U, tedy

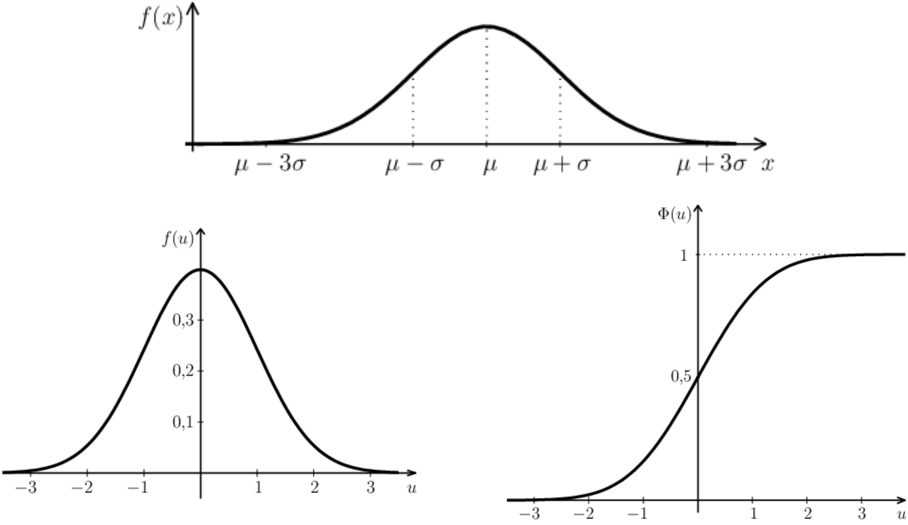
$$U \sim N(0, 1).$$

Hustota ϕ a distribuční funkce Φ náhodné veličiny U jsou

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{u^2}{2}}, \quad u \in \mathbb{R},$$

$$\Phi(u) = P(U \leq u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{t^2}{2}} \, \mathrm{d}t, \quad u \in \mathbb{R}.$$

Hodnoty funkce Φ lze najít v tabulkách.



f(x)
μ − 3σ
μ − σ
μ
μ + σ
μ + 3σ
x
f(u)
0,3
0,2
0,1
−3
−2
−1
1
2
3
u
Φ(u)
1
0,5
−3
−2
−1
1
2
3
u

Nechť X je náhodná veličina s normálním rozdělením se střední hodnotou μ a rozptylem σ^2, tedy

$$X \sim N(\mu, \sigma^2).$$

Pak

$$U = \frac{X - \mu}{\sigma}$$

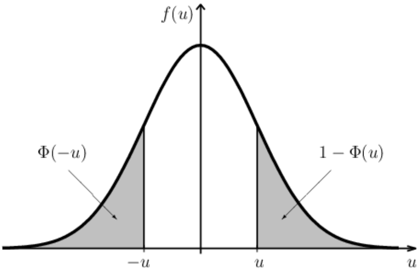
je náhodná veličina se standardizovaným normálním rozdělením, tedy

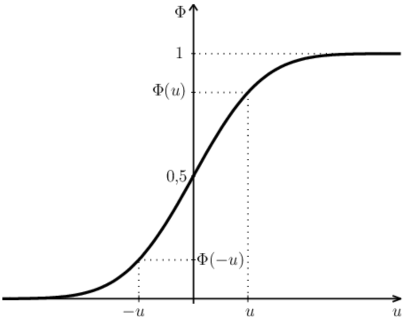
$$U \sim N(0, 1).$$

Platí:

$$P(X \leq x) = P(\mu + \sigma U \leq x) = P\left(U \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

$$\Phi(-u) = 1 - \Phi(u)$$





Příklad

Náhodná veličina X udávající spotřebu paliva naftové elektrocentrály během 8hodinové směny má normální rozdělení se střední hodnotou 9 l a rozptylem 0,16 l^2. Určete

a) pravděpodobnost, že spotřeba bude větší než 9,5 l,

b) pravděpodobnost, že spotřeba bude mezi 8,6 a 9,3 l,

c) pod jakou hranicí leží spotřeba s pravděpodobností 0,99.

Řešení:

X...spotřeba nafty

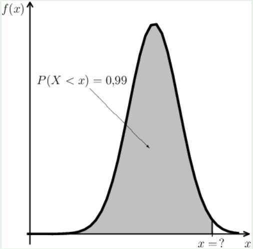
$$X \sim N(9; 0{,}16) \quad \Rightarrow \quad U = \tfrac{X-9}{0{,}4} \sim N(0,1)$$

$$P(X > 9{,}5) = P(U > 1{,}25) = 1 - P(U \leq 1{,}25) = 1 - \Phi(1{,}25) \doteq 0{,}106$$

$$P(8{,}6 \leq X \leq 9{,}3) = P(-1 \leq U \leq 0{,}75) = \Phi(0{,}75) - \Phi(-1) =$$

$$= \Phi(0{,}75) - 1 + \Phi(1) \doteq 0{,}615$$

Pod jakou hranicí leží spotřeba s pravděpodobností 0,99:



$$P(X < x) = 0{,}99$$

$$P\left(U < \frac{x-9}{0{,}4}\right) = 0{,}99$$

$$\Phi\left(\frac{x-9}{0{,}4}\right) = 0{,}99 \qquad \Phi(2{,}33) \doteq 0{,}99 \quad \text{(z tabulek)}$$

$$\frac{x-9}{0{,}4} \doteq 2{,}33 \quad \Rightarrow \quad x \doteq 9{,}932$$

Rozdělení součtu a průměru

Jestliže $X_1, X_2, \dots, X_n$ jsou navzájem nezávislé náhodné veličiny s rozdělením $N(\mu, \sigma^2)$, pak pro součet a průměr těchto náhodných veličin

$$Y = \sum_{i=1}^{n} X_i \quad \text{a} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

platí

$$Y \sim N\left(n\mu, n\sigma^2\right) \quad \text{a} \quad \overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Centrální limitní věta

Jestliže $X_1, X_2, \ldots, X_n$ jsou navzájem nezávislé náhodné veličiny, které mají všechny stejné rozdělení se střední hodnotou μ a konečným rozptylem σ^2, pak pro součet a průměr těchto náhodných veličin

$$Y = \sum_{i=1}^{n} X_i \quad \text{a} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

platí

$$\lim_{n\to\infty} P\left(\frac{Y - n\mu}{\sqrt{n}\sigma} \leq u\right) = \Phi(u), \qquad Y \overset{A}{\sim} N\left(n\mu, n\sigma^2\right)$$

$$\lim_{n\to\infty} P\left(\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq u\right) = \Phi(u), \qquad \overline{X} \overset{A}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

pro každé $u \in \mathbb{R}$.

Příklad

Životnost určitého typu součástek (měřená ve stovkách hodin) je náhodná veličina X s hustotou

$$f(x) = \begin{cases} \frac{x}{2} \cdot e^{-\frac{x^2}{4}} & \text{pro } x > 0, \\ 0 & \text{pro } x \leq 0, \end{cases}$$

střední hodnotou $\sqrt{\pi}$ a rozptylem $4 - \pi$. Náhodně vybereme 50 součástek.

a) Jaká je pravděpodobnost, že průměrná životnost těchto 50 součástek bude vyšší než 170 hodin?

b) Určete interval souměrný kolem střední hodnoty, ve kterém bude průměrná životnost těchto 50 součástek s pravděpodobností 0,95.

Řešení:

X...životnost, $X \sim ?$

$\overline{X}$...průmerná životnost, $\overline{X} \sim ?$ (CLV) $\overline{X} \overset{A}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$

$\overline{X} = \frac{1}{50}\sum_{i=1}^{50} X_i$, $\overline{X} \overset{A}{\sim} N\left(\sqrt{\pi}, \frac{4-\pi}{50}\right)$

a) $$P(\overline{X} > 1{,}7) \doteq P\left(U > \frac{1{,}7 - \sqrt{\pi}}{\sqrt{\frac{4-\pi}{50}}}\right) \doteq P(U > -0{,}55) \doteq 0{,}709$$

b) Interval souměrný kolem střední hodnoty:

$$P(x_1 < \overline{X} < x_2) = 0{,}95$$

$$P(\overline{X} < x_2) = 0{,}975$$

$$P\left(U < \frac{x_2 - \sqrt{\pi}}{\sqrt{\frac{4-\pi}{50}}}\right) \doteq 0{,}975$$

$$P\left(U < \frac{x_2 - \sqrt{\pi}}{\sqrt{\frac{4-\pi}{50}}}\right) \doteq 0{,}975$$

$$\Phi\left(\frac{x_2 - \sqrt{\pi}}{\sqrt{\frac{4-\pi}{50}}}\right) \doteq 0{,}975 \qquad \Phi(1{,}96) \doteq 0{,}975 \quad \text{(z tabulek)}$$

$$\frac{x_2 - \sqrt{\pi}}{\sqrt{\frac{4-\pi}{50}}} \doteq 1{,}96$$

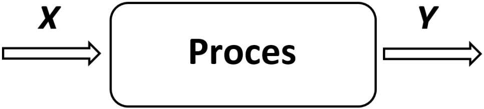
$$x_2 \doteq \sqrt{\pi} + 1{,}96\sqrt{\frac{4-\pi}{50}} \doteq 2{,}029$$

$$x_1 \doteq \sqrt{\pi} - 1{,}96\sqrt{\frac{4-\pi}{50}} \doteq 1{,}516$$

Transformace náhodných veličin

- Popis reálného procesů - výstupní hodnota y závisí na vstupní hodnotě x.
- Vstupní veličina má náhodný charakter $\Rightarrow$ výstupní veličina má náhodný charakter.

$$Y = h(X)$$



- Ze známého rozdělení vstupní veličiny určujeme rozdělení výstupní veličiny.
- Už jsme viděli v definici momentů (X^2, X^k,...), u normálního rozdělení.

Transformace diskrétní náhodné veličiny

Nechť X je náhodná veličina a h je reálná funkce taková, že $H(X) \subseteq D(h)$. Pak náhodnou veličinu Y definovanou vztahem

$$Y = h(X)$$

nazveme **transformovanou náhodnou veličinou.**

Mějme diskrétní náhodnou veličinu X s pravděpodobnostní funkcí p_X. Pak transformovaná náhodná veličina $Y = h(X)$ je také diskrétní a pro její pravděpodobnostní funkci p_Y platí

$$p_Y(y) = \sum_{x \in h^{-1}(y)} p_X(x), \quad y \in \mathbb{R}.$$

Příklad

Náhodná veličina X má pravděpodobnostní funkci

$$p_X(x) = \begin{cases} \frac{1}{5} & \text{pro } x \in \{-2, -1, 0, 1, 2\}, \\ 0 & \textit{jinak}. \end{cases}$$

Určete $p_Y(y)$, jestliže $Y = |X|$.

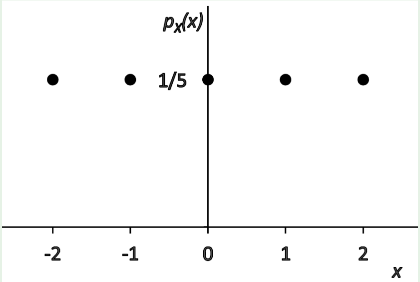
Řešení:

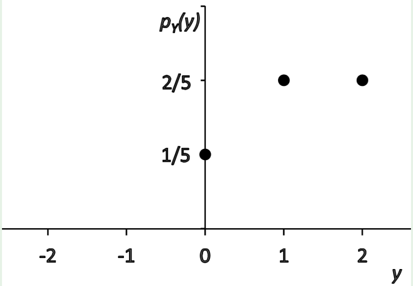
$Y = |X|$

x	-2	-1	0	1	2
$p_X(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

x	-2	-1	0	1	2
$p_X(x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
$y = \|x\|$	2	1	0	1	2

y	0	1	2
$p_Y(y)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$





Transformace spojité náhodné veličiny

Mějme spojitou náhodnou veličinu X s hustotou $f_X(x)$ a distribuční funkcí $F_X(x)$. Pak pro transformovanou náhodnou veličinu $Y = h(X)$ s hustotou $f_Y(y)$ a distribuční funkcí $F_Y(y)$ platí

$$F_Y(y) = P(Y \leq y) = P(h(X) \leq y) = P(X \in h^{-1}((-\infty, y\rangle)), \quad y \in \mathbb{R},$$

kde $h^{-1}((-\infty, y\rangle)$ značí vzor množiny $(-\infty, y\rangle$.

Příklad

Náhodná veličina X má rovnoměrné rozdělení $Ro(0, 1)$. Určete hustotu a distribuční funkci náhodné veličiny $Y = \sqrt{X}$.

Řešení:

$X \sim Ro(a, b)$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{pro } x \in \langle a, b\rangle\,, \\ 0 & \text{jinak,} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{pro } x < a, \\ \frac{x-a}{b-a} & \text{pro } x \in \langle a, b\rangle\,, \\ 1 & \text{pro } x > b. \end{cases}$$

Řešení:

$X \sim Ro(0,1), \quad Y = \sqrt{X}$

$$f_X(x) = \begin{cases} 1 & \text{pro } x \in \langle 0,1 \rangle \\ 0 & \text{jinak} \end{cases} \qquad F_X(x) = \begin{cases} 0 & \text{pro } x < 0 \\ x & \text{pro } x \in \langle 0,1 \rangle \\ 1 & \text{pro } x > 1 \end{cases}$$

$$F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F_X(y^2)$$

Graf $y = \sqrt{x}$: Vidíme, že $f_Y(y)$ je nenulová pro $y \in \langle 0,1 \rangle$.

$$F_Y(y) = \begin{cases} 0 & \text{pro } y < 0 \\ y^2 & \text{pro } y \in \langle 0,1 \rangle \\ 1 & \text{pro } y > 1 \end{cases}$$

$$f_Y(y) = F'_Y(y) = \begin{cases} 2y & \text{pro } y \in \langle 0,1 \rangle \\ 0 & \text{jinak} \end{cases}$$

Příklad

Pepa pluje trajektem 165 km z Rostoku do Trelleborgu konstantní rychlostí V, která má rovnoměrné rozdělení mezi 50 a 90 km/h. Určete hustotu a distribuční funkci doby trvání plavby T.

Řešení:

$V \sim Ro(50, 90), \quad T = \frac{165}{V}$

$$f_V(v) = \begin{cases} \frac{1}{40} & \text{pro } v \in \langle 50, 90 \rangle \\ 0 & \text{jinak} \end{cases} \qquad F_V(v) = \begin{cases} 0 & \text{pro } v < 50 \\ \frac{v-50}{40} & \text{pro } v \in \langle 50, 90 \rangle \\ 1 & \text{pro } v > 90 \end{cases}$$

$$F_T(t) = P(T \le t) = P\left(\frac{165}{V} \le t\right) = P\left(\frac{165}{t} \le V\right)$$

$$= 1 - P\left(V < \frac{165}{t}\right) = 1 - F_V\left(\frac{165}{t}\right)$$

$$F_T(t) = 1 - F_V\left(\frac{165}{t}\right), \qquad F_V(v) = \begin{cases} 0 & \text{pro } v < 50 \\ \frac{v-50}{40} & \text{pro } v \in \langle 50, 90\rangle \\ 1 & \text{pro } v > 90 \end{cases}$$

Graf $t = \frac{165}{v}$: Vidíme, že $f_T(t)$ je nenulová pro $t \in \left\langle \frac{165}{90}, \frac{165}{50} \right\rangle$.

$$F_T(t) = 1 - \begin{cases} 0 & \text{pro } \frac{165}{t} < 50 \\ \frac{\frac{165}{t}-50}{40} & \text{pro } \frac{165}{t} \in \langle 50, 90\rangle \\ 1 & \text{pro } \frac{165}{t} > 90 \end{cases} = \begin{cases} 1 & \text{pro } \frac{165}{50} < t \\ \frac{9}{4} - \frac{165}{40t} & \text{pro } t \in \left\langle \frac{165}{90}, \frac{165}{50} \right\rangle \\ 0 & \text{pro } \frac{165}{90} > t \end{cases}$$

$$= \begin{cases} 0 & \text{pro } t < \frac{11}{6} \\ \frac{9}{4} - \frac{33}{8t} & \text{pro } t \in \left\langle \frac{11}{6}, \frac{33}{10} \right\rangle \\ 1 & \text{pro } t > \frac{33}{10} \end{cases}$$

$$f_T(t) = F_T'(t) = \begin{cases} \frac{33}{8t^2} & \text{pro } t \in \left\langle \frac{11}{6}, \frac{33}{10} \right\rangle \\ 0 & \text{jinak} \end{cases}$$

Příklad

Náhodná veličina X má normální rozdělení $N(0,1)$. Určete hustotu a distribuční funkci náhodné veličiny $Y = X^2$.

Řešení:

$X \sim N(0,1)$

$\phi(x) = \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{x^2}{2}},\ x \in \mathbb{R}, \qquad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t, \quad x \in \mathbb{R}$

$$F_Y(y) = P(Y \le y) = P\left(X^2 \le y\right) = P\left(-\sqrt{y} \le X \le \sqrt{y}\right)$$

$$= \Phi\left(\sqrt{y}\right) - \Phi\left(-\sqrt{y}\right) = 2\Phi\left(\sqrt{y}\right) - 1, \quad y \ge 0$$

$$f_Y(y) = F_Y'(y) = \frac{\phi\left(\sqrt{y}\right)}{\sqrt{y}} = \frac{1}{\sqrt{2\pi y}}\,\mathrm{e}^{-\frac{y}{2}}, \quad y \ge 0, \qquad \text{rozdělení } \chi^2(1)$$

Mějme spojitou náhodnou veličinu X s hustotou $f_X(x)$ a distribuční funkcí $F_X(x)$. Dále nechť h je ryze monotónní funkce, která má všude derivaci. Pak transformovaná náhodná veličina $Y = h(X)$ má hustotu

$$f_Y(y) = f_X\left(h^{-1}(y)\right)\left|\frac{\mathrm{d}h^{-1}(y)}{\mathrm{d}y}\right|,$$

kde h^{-1} značí inverzní funkci k funkci h.

Příklad

X, $F_X(x)$, $f_X(x)$

$$Y = \underbrace{\sqrt{X}}_{h}, \quad X = \underbrace{Y^2}_{h^{-1}}$$

$$F_Y(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F_X(y^2)$$

$$f_Y(y) = F'_Y(y) = \left[F_X(y^2)\right]' = f_X(\underbrace{y^2}_{h^{-1}}) \cdot \underbrace{2y}_{\left(h^{-1}\right)'}$$

Střední hodnota transformované náhodné veličiny

Střední hodnota transformované náhodné veličiny $Y = h(X)$ je (v případě, že existuje) definována jako

$$\mathrm{E}(Y) = \mathrm{E}(h(X)) = \sum_{x \in M} h(x) \cdot p_X(x)$$

pro rozdělení diskrétního typu s oborem hodnot M a

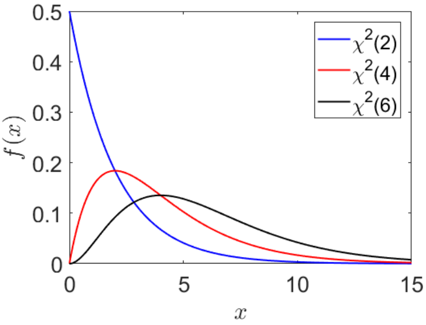
$$\mathrm{E}(Y) = \mathrm{E}(h(X)) = \int_{-\infty}^{\infty} h(x) \cdot f_X(x) \, \mathrm{d}x$$

pro rozdělení spojitého typu.

Nechť $U_1, \ldots, U_n$ jsou nezávislé náhodné veličiny s rozdělením $N(0,1)$. Pak

$$K = \sum_{i=1}^{n} U_i^2 \sim \chi^2(n)$$

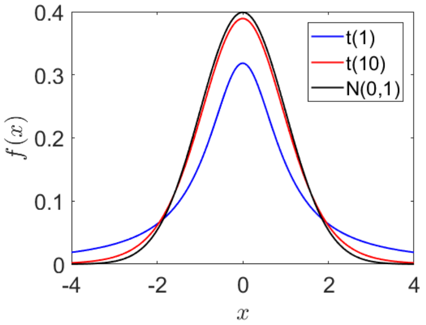
má χ^2 rozdělení o n stupních volnosti.



Mějme dvě nezávislé náhodné veličiny $U \sim N(0,1)$ a $K \sim \chi^2(n)$. Pak

$$T = \frac{U}{\sqrt{\frac{K}{n}}} \sim t(n)$$

má Studentovo t rozdělení o n stupních volnosti.



Mějme dvě nezávislé náhodné veličiny $K_1 \sim \chi^2(m)$ a $K_2 \sim \chi^2(n)$. Pak

$$F = \frac{\frac{K_1}{m}}{\frac{K_2}{n}} \sim F(m, n)$$

má Fisherovo-Snedecorovo F rozdělení o m a n stupních volnosti.